Editor’s Message: Teetering Atop the Pyramid: What to Do About Systematic Reviews?

A systematic review [SR] is a secondary research study that uses defined, reproducible and transparent methods to locate, screen, and critically assess all previously published primary research studies designed to answer an important clinical question. From the articles included in a SR, pertinent data is described, summarized and analyzed. While SRs may cover observational studies, or address measurement topics, here I refer primarily to SRs of clinical intervention trials. About 1/3 of SRs include a meta-analysis [SRMA] in which statistical techniques are used to synthesize data from several similar studies in a single, summary quantitative estimate of the effect size of the studied intervention compared to either no intervention or a comparison intervention. The result of the meta-analysis of multiple studies is generally considered to be stronger evidence than the result of any single study, although this is not always the case. Overall, the publication rate for SR/SRMAs in healthcare journals has increased dramatically, outstripping the publication rate for original clinical trials. This trend is concerning for reasons described elsewhere, however, this trend is not seen rehabilitation journals, where (at least until recently) clinical trials continue to be published at a higher rate than SRs.

Systematic reviews are considered the highest level of evidence, and for this reason they are trusted sources of information to guide clinical decisions. They are popular with researchers, clinicians, and journals alike. Researchers investigating related topics value having a thorough sense of what is known and what is not known (knowledge gap) about a topic prior to designing and conducting their own studies. Clinicians striving for evidence-based practice who struggle to stay abreast of hundreds of published studies appreciate that an expert has done the leg-work and produced a useful summary for clinical decision-making. Journals seeking to improve their impact factor like the fact that SRs are frequently cited. What’s not to like, right? As it turns out, plenty.

While there is little doubt that the benefits of high-quality SRs are many, there is widespread consensus that poor quality SRs are at best misleading and thus potentially result in erroneous clinical decisions. Further, with the caveat that the quality of more recent SRs is increasing, extensive reviews of published SRs have found that a shockingly large majority are of low quality. Real and consequential problems with SRs have been recognized and discussed, sometimes heatedly, since the 1990s. These problems include, but are not limited to: redundancy, unimportant or overly broad research questions, insufficiently rigorous (and sometimes seriously biased or even deceitful) conduct of the review process, the inclusion of primary studies that suffer from bias and/or were not conducted well themselves, journal publication bias toward studies with significant results, substantial differences between included studies, and the cobbling together of fragmented evidence in order to produce an apparently seamless whole conclusion. A number of highly respected researchers have come to doubt the usefulness of SRs, with several arguing forcefully that SRs do more harm than good and should no longer be published.

Other researchers counter that the quality of SRs has improved markedly, and these efforts to synthesize evidence for clinicians should not be abandoned. Many steps have been taken to improve the quality of SRs. Prospective registration of SRs is now required through PROSPERO or other registration sites. Research questions are guided by the PICO framework (Population, Intervention, Comparison, Outcome). Standardized guidelines for the procedural conduct of the SR study must be followed (e.g., PRISMA). Multiple authors are required, at least two who independently rate the candidate articles for inclusion and extract the data, with an additional author to resolve any discrepancies. Standardized tools to assess the quality of included articles (e.g., GRADE) and their risk of bias (e.g., Cochrane Risk of Bias – 2) are available and used for responsible reporting of the limitations of those included studies. Researchers supportive of the continued publication of SRs do acknowledge the many inherent problems. However, they posit that the need for evidence synthesis to support evidence-based clinical practice is real and growing, and that currently, no better alternative to SRs exists. These same authors pull no punches when it comes to SR quality, however, stating outright that editors should not publish un-registered or low-quality SRs.

Based on my review of numerous SR submissions to the Journal of Geriatric Physical Therapy, the largest problem with SRs in the rehabilitation field is the sub-par quality of previously published clinical trials, a great number of which would not be accepted for publication by today’s standards. Too many have a very high risk of bias because there was no randomization, no assessor blinding, or indeed, no control group at all. The use of invalid, unreliable or unresponsive outcome measures often compounds the problems of interventions that do not specifically
target the clinical problem (mode) and/or are not delivered with sufficient frequency, intensity, duration, or progression to produce any expected change. Failure to employ multivariate statistical analyses with correction for multiple comparisons results in the high probability of Type I error and reported statistical significance where none existed. Inadequate sample sizes produce under-powered studies with a high probability of Type II error and reported non-significance when significance may have existed had the study been powered to find it. Traditional results reporting did not present important measures of clinical significance, e.g., 95% confidence intervals, effect sizes, and whether or not post-intervention changes met or exceeded the minimal detectable change for each outcome measure. The limitations of many of these original studies negatively impact the validity of their results and conclusions; when several excessively limited studies are included in a SR, the validity of the SR results and conclusions are themselves negatively impacted. Per Chevret and colleagues, this is the “garbage in, garbage out” problem. Study registration and standardized guidelines for SR procedures do not address this problem. Standardized approaches to rating and reporting study quality and risk of bias may identify these problems, but do not solve them. The serious limitations of many previously published intervention trials are not the fault of the SR authors, but if SR authors choose to include them, the conclusions of the SR are likely to fall along the lines of “inadequate evidence, inconclusive results, more studies are needed ...”. These SR conclusions may be accurate, and may point researchers to the great need for high quality intervention studies of that topic, but they do not offer any support for clinical decision making. This particular problem will not be resolved until many more high-quality clinical rehabilitation trials are published, then subsequently included in SRs.

Journalist and writer Eric Sevareid said: “The chief cause of problems is solutions.” Systematic reviews are today the best solution we have developed thus far to address the need for evidence collection and synthesis to support evidence-based practice. They are an admittedly imperfect solution, with many problems leading to a proliferation of poor quality SRs. The Editorial Team of the JGPT is aware of these potential problems, and aims to publish only high-quality SRs. We now ask authors who submit SR manuscripts to exclude original studies that are clearly of low quality or high bias from their final analysis, and to exert extra effort to document that the quality of their included studies is at least adequate. As the number of properly designed and conducted clinical intervention trials increases, and these publications are subsequently included in SRs, the quality of rehabilitation SRs will also rise.

REFERENCES
3. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. Milbank Q. 2016;94(3):485-514.